JOURNAL OF TIME SERIES ANALYSIS
J. Time Ser. Anal. 45:333-360 (2024)
Published online 25 August 2023 in Wiley Online Library
(wileyonlinelibrary.com) DOI: I0.Ill l/jtsa.12714

ORIGINAL ARTICLE

STATIONARY JACKKNIFE

WEILIAN ZHou·· ANDSOUMENDRA LAHIRibG)
'Department of Statistics, North Carolina State University, Raleigh, NC, USA
[b] *Department of Mathematics and Statistics, Washington University, St. Louis, MO, USA*

Variance estimation is an important aspect **in** statistical inference, especially in the dependent data situations. Resampling methods are ideal for solving this problem since these do not require restrictive distributional assumptions. In this paper, we develop a novel resampling method in the Jackknife family called thestationary jackknife. It can be used to estimate the variance of a statistic in the cases where observations are from a general stationary sequence. Unlike the moving block jackknife, the stationary jackknife computes the jackknife replication by deleting a variable length block and the length bas a truncated geometric distribution. Under appropriate assumptions, we can show thestationary jackknife variance estimator is a consistent estimator for the case of the sample mean and, more generally, for a class of nonlinear statistics. Further, the stationary jackknife isshown to provide reasonable variance estimation for a wider range of expected block lengths when compared with the moving block jackknife by simulation.

Received 13 June 2022; Accepted 29 June 2023

Keywords: Stationary jackknife; variance estimation; consistency; moving block jackknife.

MOS subject classification: 62F40; 62M12.

1. INTRODUCTION

The Jackknife (cf. Quenouille, 1949, 1956; Tukey, 1958) is an intriguing non-parametric method for estimating the bias and variance of statistics we are interested in. The role of the jackknife in bias correction and robust confidence interval has been fully explained in Miller (1974). A general resampling method, the bootstrap method (cf. Efron, 1979) is introduced to work satisfactorily on a variety of estimation problems. Efron and Gong (1983) provide a comparison between the jackknife, the bootstrap and another important non-parametric method, cross validation at an accessible mathematical level. However, for all these methods, the assumption of the independence of the observations is very crucial. It seems that the standard jackknife or bootstrap will give an unreliable estimation if dependence is ignored. In most cases, especially in applications involving time series, dependence between observations is not negligible. When fitting a parametric model to a given time series, it is always very difficult to model all important features of the observed time series, and the parametric inference approach often suffers from the risk brought in by the effect of the parameter estimation or model misspecification. Thus, it is very important to modify the jackknife estimator for the dependent data.

Carlstein (1986) proposed a block-wise resampling method where the variance estimator is computed using non-overlapping blocks. In some cases, especially for the arithmetic mean, selection of blocks is equivalent to the deletion of complementary blocks. The moving block jackknife (bootstrap) of Kunsch (1989), an extension of the standard jackknife (bootstrap) method, computes the statistics of interest by deleting (selecting) overlapping fixed length data blocks. It shows that the moving block method works well for arbitrary stationary time series with short range dependence. The circular block bootstrap of Politis and Romano (1992) wraps the data in a

• correspondence to: Weilian Zhou, Department of Statistics, North Carolina State University, Raleigh, NC 27606, USA.
Email: wzhoul l@ncsu.edu

circle before selecting the overlapping blocks. Every data point is assigned the same probability mass in the conditional distribution (condition on the original data) by the circular block bootstrap. This provides a more accurate approximation with respect to the mean. The stationary bootstrap of Politis and Romano (1994) selects blocks of random lengths to make the bootstrap observations stationary. Some theoretical results about the stationary bootstrap have been obtained in Lahiri (1999) and Nordman (2009). It is surprising that the variance of the stationary bootstrap matches that of a block bootstrap based on non-random, non-overlapping blocks. For more details on related properties of block bootstrap methods, see Lahiri (2003).Some other developments in the jackknife resampling literature, such as the delete-d jackknife (Wu, 1986; Shao and Wu, 1989), the threshold jackknife (Park and Willemain, 1999) and the artificial jackknife (Pellegrino, 2022), discuss the variance estimation in dependent data situations. In this article, we introduce a new resampling method called the stationary jackknife. Similar to the extension made in the stationary bootstrap (Politis and Romano, 1994) on the moving block bootstrap, our method is anextension of the moving block jackknife. We apply a variable block length in each deleted block. This is the reason why we call our method the stationary jackknife.

The stationary jackknife is suitable for variance estimation of statistics from observations generated by weakly dependent stationary time series.The stationary jackknife can be applied insituationssimilar to the moving block jackknife. The major difference between the stationary jackknife and the moving block jackknife lies in the block length. For the stationary jackknife, we delete *l;* consecutive observations as the *ith* block to compute the *ith* pseudo value. In the settings of the moving block jackknife, *t',* = *t',* the block length is a constant for all pseudo-value computation. However in the stationary jackknife, *t';*isnot necessarily the same every time and we treat *t',* as a random variable with a truncated geometric distribution. Since a variable with the geometric distribution is unbounded, we set an upper bound on the variable to avoid deleting a large segment of the observations. Moreover, the fact that the tail of the geometric distribution decays exponentially helps constrain the block length to within a reasonable range. For the starting point of the block near the end of the observations, the last observation serves as a natural cutting point for the block. After deleting the block, we perform a smooth transformation on the remaining observations to get the statistic. The stationary jackknife variance estimator is the standardized version of sample variance of statistics obtained in the way mentioned above. One common difficulty in the moving block jackknife is the choice of the optimal order of the block length. The optimal order of the block length may vary when jackknifing different statistics or thesame class of statistics under different data generatingdistributions. Nonetheless, we observed that the randomness in the block length makes the stationary jackknife method more robust to the expected block length. The simulation results in Section 5 illustrate that the stationary jackknife provides a reasonable variance estimation in a wider range of the expected block length compared with the moving block jackknife.

In Section 2, we introduce the stationary jackknife and illustrate its differences with the moving block jackknife. In Section 3, we prove the consistency of the stationary jackknife variance estimator in the case of the arithmetic mean. We derive the asymptotic bias and variance terms of the stationary jackknife variance estimator. By minimizing the mean square error (MSE), we get the optimal order of the expected block length for the stationary jackknife. In Section 4, we investigate the consistency of the stationary jackknife variance estimator for more general statistics. In Section 5, we compare thestationary jackknife with different resampling methods suchas the movingblockbootstrapin Kunsch (1989), the moving block jackknife and the stationary bootstrap in Politis and Romano (1994) on different simulated datasets.

2. FORMULATION OFTHE STATIONARY JACKKNIFE METHOD

We formalize the definition of the stationary jackknife. One important issue is that we need to define the statistic with a variable length missing block of the observations. The issue can be overcome by focusing on a certain class of statistics we are going to introduce later in this section. This class of statistics is sufficiently rich to include many commonly used statistics.

DOI: 10.1111/jtsa.12714

2.1. Estimator defined on the empirical distribution

For observations $\{X_1, X_2, \dots, X_N\}$ from a stationary process, following the definition in section 2.1 of Kunsch (1989), we define the empirical k-dimensional marginal distribution as

$$\rho_N^k = (N-k+1)^{-1}\sum_{i=1}^{N-k+1}\delta_{(X_i, X_{i+1}, \dots, X_{i+k-1})},$$

where δ_y is the point mass on $y \in \mathbb{R}^k$. For any functional T defined on all probability measures on $\mathbb{R}^k$, we consider the statistic T_N of the form

$$T_N\left(X_1, X_2, \dots, X_N\right) = T\left(\rho_N^k\right).$$

Denote $n = N - k + 1$ as the number of k-tuples in the observations. For notational simplicity, we use the equivalent relationship below as in the examples in the section 2.1 of Kunsch (1989), by setting

$$\rho_N^k \equiv \rho_n,$$

where ρ_n is the empirical distribution of n observations.

2.2. The stationary jackknife

Before deriving the formula for the stationary jackknife, we recall how the moving block jackknife works to produce the pseudo values. In the moving block jackknife, the length of the block ℓ is fixed over the time. In producing the jth pseudo value, the corresponding marginal $\tilde{\rho}_n^{(j)}$ after deleting the jth data block as defined in Kunsch (1989) is,

$$\tilde{\rho}_n^{(j)} = \frac{1}{n - \|w_n\|_1}\sum_{t=1}^{n}\left(1 - w_n(t-j)\right)\delta_{Y_t}, \quad j = 1, 2, \dots, n - \ell + 1,$$

where w_n represents the scale of downweight for each observation, and w_n satisfies the following properties

$$\begin{aligned} 1 \ge w_n(i) &> 0 \quad 0 \le i < \ell, \\ w_n(i) &= 0 \quad \text{otherwise,} \end{aligned}$$

and $\|w_n\|_1 = \sum_i w_n(i)$. In our case, we focus on removing the data blocks rather than downgrading them. So we have $w_n(i) = \mathbb{1}_{([0, \ell-1])}(i)$ and the corresponding marginal $\tilde{\rho}_n^{(j)}$ can be written as

$$\tilde{\rho}_n^{(j)} = \frac{1}{n-\ell}\sum_{t=1}^{n}\left(1 - \mathbb{1}_{([j, j+\ell-1])}(t)\right)\delta_{Y_t}, \quad j = 1, 2, \dots, n - \ell + 1.$$

Then the jth moving block jackknife replication is represented as the estimator defined on the empirical marginal $\tilde{\rho}_n^{(j)}$,

$$\tilde{T}_n^{(j)} = T\left(\tilde{\rho}_n^{(j)}\right).$$

DOI: 10.1111/jtsa.12714

Finally, the moving block jackknife variance estimator $\hat\sigma^2_{Jack}$ is defined as

$$\hat\sigma^2_{Jack} = \frac{n-\ell}{n\ell}\,\frac{1}{n-\ell+1}\sum_{i=1}^{n-\ell+1}\left(\tilde T_n^{(i)} - \tilde T_n\right)^2, \tag{1}$$

where $\tilde T_n(\cdot) = \frac{1}{n-\ell+1}\sum_{i=1}^{n-\ell+1} \tilde T_n^{(i)}$.

For the stationary jackknife, the deleting block length L follows a truncated geometric distribution $TGeo(p, T)$ with the probability density function (pdf) as following.

$$P(L=k) = \begin{cases} p(1-p)^{k-1} & 1 \le k \le T-1, \\ (1-p)^{[\ell \log n]-1} & k = T, \end{cases}$$

where T is the truncated value and p is the parameter of the geometric distribution. In the stationary jackknife, we set $T = [2\ell \log n]$ (with $\ell = \frac{1}{p}$ and $[a]$ representing the largest integer not exceeding a). The corresponding cumulative density function F_L is

$$F_L(k) = \begin{cases} 0 & k < 1, \\ 1-(1-p)^k & 1 \le k \le T-1, \\ 1 & k \ge T. \end{cases}$$

As we can see, the main difference between the stationary jackknife and the moving block jackknife is that the length of our deleted block is no longer a constant but a random variable. We set an upper bound $[2\ell \log n]$ on the block length to make sure the length of the missing part is much smaller than the length of the original observations.

To derive the expression of the stationary jackknife, we need to define $\{L_1, L_2, \ldots, L_m; m = n - [2\ell \log n] + 1\}$ to represent the block length variables. More specifically, $\{L_1, L_2, \ldots, L_m\}$ are independent and each has the same distribution as L. The realization of $\{L_i\}$ can be generated by a series of geometric distribution variable $\{\tilde L_i\}$ due to the following relationship

$$L = \min(\tilde L, [2\ell \log n]),$$

where $\tilde L$ is a random variable with a geometric distribution $\tilde L \sim Geo(p)$. In the jth pseudo value computation, the corresponding empirical marginal $\tilde P_n^{(j)}$ is defined as

$$\tilde P_n^{(j)} = \frac{1}{n - L_j}\sum_{t=1}^{n}\left(1 - \mathbb{1}([j, j+L_j-1])(t)\right)\delta_{X_t}, \quad j = 1, 2, \ldots, n - [2\ell \log n] + 1.$$

Then the jth stationary jackknife replication is calculated as

$$\tilde T_n^{(j)} = T(\tilde P_n^{(j)}), \quad j = 1, 2, \ldots, n - [2\ell \log n] + 1.$$

From the definition of the stationary jackknife replication, there are two layers of randomness in the stationary jackknife, one comes from the empirical distribution, the other one comes from the realizations of the truncated geometric distribution.

DOI: 10.1111/jtsa.12714

The stationary jackknife estimator of the variance of *Tn* is a standardized version of the sample variance of the *T';/1*

$$\hat{\sigma}^2_{\text{jack}} = \frac{(n-\ell)^2}{n\ell m}\sum_{j=1}^{m}\left(T_n^{(j)} - T_n^{(\cdot)}\right)^2, \quad (2)$$

where *m* = *n* - [2t'logn] + 1 and r-·) n =r:;,r,:'.m The estimator can be decomposed into two parts. The first part is sample variance of the stationary jackknife replications ; r; I (T';P - -))^2. The second part is the standardizing factor (n - ..!) n e ^{2}. The standardizing factor is inversely proportional to the expected block length *t'*. The larger the expected deleting length we choose, the larger the ratio we down weight on the sample variance of the jackknife replications.Thestandardizing factor in formula (2) is thesame as thestandardizing factor of the moving jackknife estimator (1). lnthe moving block jackknife, *n* - *t'* can be treated as the remaining observations in each jackknife replication, and in the stationary jackknife, *n* - *t'* can be treated as the expected remaining observation in each replication. Moreover, the standardizing factor helps the estimator enjoy the consistency property reported in Section 3.

3. CONSISTENCY OF THE STATIONARY JACKKNIFE ESTIMATOR

Here, we show the consistency of the stationary jackknife variance estimator for the arithmetic mean under several assumptions. The assumptions are similar to the assumptions for the moving block jackknife. The routine method inproving theconsistency of the estimator is toshow both the bias and the variance of theestimator converge to zero. We derive the exact formula of the bias and variance in the following theorems. Moreover, we investigate the optimal expected block length by minimizing the mean squared error of the estimator.

Here, we investigate the properties of the stationary jackknife for the arithmetic mean. The arithmetic mean is obviously a member of the class of estimators defined in section 2.1. It corresponds to the population mean *T (F)* = *f xF (dx)* = lE*(x)* = *μ*. The functional is linear and allows for explicit calculations of all quantities of interest.

Suppose that {X1}r-t,z,... oo are from a weakly stationary processes. Then *Xl* enjoys the following properties,

$$\text{lE}(X,) = \mu.$$
$$\text{lE}(X;) < oo.$$
$$\text{IE}((Xt+h - \mu)(X, - \mu)) = R(h), \quad Vt.$$

According to the stationarity of the observations, we can write the the standardized variance of the arithmetic mean, var(*ynX*), analytically as

$$\text{Var}\left(\sqrt{n}\bar{X}\right) = \sum_{h=-n+1}^{n-1}\frac{(n-|h|)}{n}R(h). \quad (3)$$

We denote *u;s* as the limit of *Var* (*ynX*). When *R(h)* is absolutely summable, then we have

$$u! = \sum_{l>=-oo}^{h-oo} R(h). \quad (4)$$

DOI: 10.I I I l/jtsa.12714

For the jth stationary jackknife replication $\hat{T}^{(j)}_n$ in the arithmetic mean case, we have

$$\hat{T}^{(j)}_n = \frac{n\bar{X}_n - S_{j,L_j}}{n - L_j},$$

where $S_{j,l} = \sum X_i$, the partial sum of $\{X_t\}$.

Before looking at the bias and the variance terms of the stationary jackknife variance estimator, we need to first specify some assumptions to ensure the consistency of the estimator. Here are some assumptions needed for the later theorems.

Assumption 1. $\ell^{-1} + \frac{\ell^2}{n} \to 0.$

Assumption 2. $\sum_{k=-\infty}^{\infty} |k| |R(k)| < \infty.$

Remark 1. Assumption 1 represents two conditions, the first one is $\frac{1}{\ell} \to 0$ and the second one is $\frac{\ell^2}{n} \to 0$. This means that the expected blocklength ℓ goes to infinity as n goes to infinity. The order of ℓ is not comparable to the observation length since we do not want to delete almost the whole series. Overall, the Assumption 1 is equivalent to

$$\ell = \ell(n) \to \infty,$$
$$\ell = o\left(n^{\frac{1}{2}}\right).$$

Assumption 2 focuses on the weak dependence of the observations. The constraint on the infinite sum of the autocovariance of $\{X_t\}$ avoids the existence of the long dependence. The autocorrelation needs to have the decay rate at least $\frac{1}{k^2}$. The exponential decay rate on the autocorrelation is a sufficient condition for the Assumption 2, so the stationary AR, MA, ARMA time series satisfy the assumption.

3.1. Bias of the stationary jackknife variance estimator

Theorem 1. Under the Assumptions 1 and 2, the bias of the stationary jackknife estimator (2) for the arithmetic mean is following.

$$\mathbb{E}\left(n\hat{\sigma}^2_{jack}\right) = \sigma^2_{as} + \mathcal{O}\left(\frac{1}{\ell}\right) + \mathcal{O}\left(\left(\frac{\ell}{n}\right)^{\frac{1}{2}}\right). \tag{5}$$

Proof See Appendix A.1. ■

To derive the variance expression of the stationary jackknife estimator $n\hat{\sigma}^2_{JDC}$, we need two additional assumptions, for some $\delta \in (0, \infty)$:

Assumption 3. $E\left(|X_1|^{6+\delta}\right) < \infty.$

Assumption 4. $\sum_{k=-\infty}^{\infty} k^2 \alpha(k)^{\frac{\delta}{6+\delta}} < \infty.$

The Assumption 3 imposes a constraint on the moment of X_1 that X_1 has a finite $(6+\delta)$th moment. Assumption 4 allows the series X_1 to be weakly dependent. It is sufficient for us to impose a strong mixing condition on X_1. For any two σ-algebras $\mathcal{A}$ and $\mathcal{B}$, define

$$\alpha(\mathcal{A}, \mathcal{B}) = \sup_{A \in \mathcal{A}, B \in \mathcal{B}} |P(A)P(B) - P(A \cap B)|. \tag{6}$$

Roughly speaking, we require the mixing coefficient *a (k)* = *a (:Ji'!} . :li't),* to decrease polynomially ask oo, where *:Ji':,²* is the a-algebra generated by {X₁}T, .

With respect to the a-mixing coefficient, we introduce a fundamental inequality for the covariance of mixing random variables.

Lemma 1. Let *X* and *Y* be measurable random variables with respect to two a-algebras *&f* and *Si.* Define IIXIIP = (IEIXI •. Then, for any *p,q,r* 1 and - + - + - = 1,

$$\text{ICov (X, Y) I} \quad 8a^{\frac{1}{7}} \; (\&f, Si) \; \text{IIXllpIIYllq} \cdot \tag{7}$$

Proof. See section 1.2.2 in Dou.khan (1994). ■

Theorem 2. Under Assumptions 1 to 4, the expression of the variance of the stationary jackknife variance estimator *n&Jack* of the arithmetic mean is

$$\text{var}(n\&^2_{\text{Jae}}k) = \frac{!}{4n} \cdot a^4_{as} + o\left(\frac{f}{n}\right). \tag{8}$$

Proof. See Appendix A.2. ■

Remark 2. Through the expression of bias and variance of the stationary jackknife estimator we can see that when *f* oo and *f* = *o (n),* both bias and variance goes to zero. This shows the consistency of our estimator in this case.

3.2. Rate of optimal block length

After obtaining the expression of bias and variance, we can obtain the optimal block length *f* in the order of *n* by minimizing the MSE of our estimator. By the standard decomposition of the MSE, we have

$$\text{MSE}(n\&J_{,.,}k) = \text{Bias}^2(n\&fack) + Var\,(n\&J_{,.,}k) = (:)(\,)^2) + (:)(\,) \cdot$$

Then we get $f = B(n^{\frac{1}{,}})$ to be the rate of optimal block length.

4. EXTENSION TO THE NONLINEAR STATISTICS

In this part, we generalize the consistency results to the nonlinear statistics. The corresponding result for the block jackknife has been well developed in Kunsch (1989). Under suitable assumptions, the stationary jackknife estimator is a consistent estimator for the nonlinear statistics.

4.1. The stationary jackknife on nonlinear statistics

The population parameter of $T_n = T(X_1, X_2, \ldots X_n)$ *is{}= T (F)* where *F* is the marginal distribution of *X.* Many estimators can be included in the nonlinear statistics. As described in Kunsch (1989), M-estimator, U-statistics (Hoeffding, 1948) and statistics defined implicitly as solutions of an equation. The linearization of T_n at F gives

$$\begin{aligned} T_n &= T(F) + n^{-1} \sum_{i=1}^{n} \text{IF}\,(X_i, F) + R_n \\ &= M_n + R_n. \end{aligned}$$

DOI: 10.1111/jtsa.12714

where IF is the influence function that IF *(x, F)=* lime o[T(O-eJF+e<>,)-T(F)]/e. Then we have

$$1:/) = T(F) + (n - Ljr' ,\_\_ L^{n}, (1 - 11.(i:Si + Lrl)) \text{ IF } (X;, F) + K\text{-}/?$$

$$= L_n^{(j)} + R_n^{(j)}.$$

For large **n,** the behavior of o}ack *(Tn)* is similar as the behavior of o}ack *(Ln)* if *R)* is in the small order. One sufficient condition for the reminder) i nsa x) E (R^2)) = *tJ* (n^{-2}).

Under this condition, it is obvious that *l',;:::)*^{2} =('JP (n-^{1}). Moreover we have

$$\sum \left(R_n^{(j)} - R_n^{(\cdot)}\right)^2 \leq \sum R_n^{(j)2},$$

and the standardized factor for the stationary jackknife estimator is *tJ* (7). Thus we have

$$n\&^2_{\text{Jack}}(T_n) = n\&^2_{\text{Jack}}(L_n) + n\&^2_{\text{Jack}}(R_n) + 2\frac{(n-t')^2}{t'm} 4 \{L_n) - L_n^{0)}) (R_n) - R_n^{0)})$$

$$n\&!ck\{Ln) + {}^2_{n\&jack}\{Rn) + 2nJ\_\overline{\&Jack (Ln) \&Jack (Rn)}$$

$$= n\&Jack (Ln) + tJP(r^{1/2}). \tag{9}$$

The reason why we get the last line in 9 is due to the following fact.

$$n\hat{\sigma}^2_{\text{jack}}(R_n) = \frac{(n-\ell)^2}{\ell m}\sum_j \left(R_n^{(j)} - R_n^{(\cdot)}\right)^2$$

$$\leq \frac{(n-\ell)^2}{\ell m}\sum_j R_n^{(j)2}$$

$$= \mathcal{O}_p\left(\frac{(n-\ell)^2}{n\ell m}\right)$$

$$= \mathcal{O}_p\left(\frac{1}{\ell}\right).$$

Together with *n&Jack* {Ln) = ('JP(1), we have $n\,\overline{\&Jack (Ln)\ \&!ck (Rn)}$ = ('JP(r½). In summary we have Theorem 3.

Theorem 3. If maxjIE (*R*)^2) = *tJ* ( ), then we have *n&Jack* (*Tn)* = no}ack (*Ln)* + *tJP* (*r* ½). Under the assumptions in Section 3, we have *n&Jack* converges to *uJs,* where

$$a = \sum_{k=-\infty}^{\infty} \text{IEIF } (X_0, F) \text{ IF } (x_k, F). \tag{10}$$

More precise we look at the linearization of Ton *Pn* instead of *F*. Assume that IF *(y, Pn)* exists, then we can write

$$T;) = T(Pn) + (n - Lj)^{-1} \sum_{i-1}^{n} P - 11.(Uj+L;-ID(L))\text{IF}(Xi, Pn) + S/,)$$
$$= M_n^{(j)} + S_n^{(j)}.$$

The distance between *p)* and *Pn* is much closer than the the distance between *p)* and *F*. The expected value of the total variation distance between *p)* and *Pn* is

$$\begin{aligned}\mathbb{E}\left(d_{\text{TV}}\left(\rho_n^{(j)}, \rho_n\right)\right) &\leq \sum_{i=1}^{n} \mathbb{E}\left|\left(1 - \mathbb{1}_{[j,j+L_j-1]}(i)\right)\frac{1}{n - \tilde{L}_j} - \frac{1}{n}\right| \\ &\leq \frac{\ell}{n} + \frac{\ell}{n - 2\ell\log n} \\ &= \mathcal{O}\left(\frac{\ell}{n}\right).\end{aligned} \tag{11}$$

From this, it is reasonable to expect S';,) to be at most of the order *('JP (* f,-). Then we have the Theorem 4

Theorem 4. If ***L*** S';,)² = *('JP (t⁴n-³)* and &jack {Mn) = *('JP* G). Then we have

$$n\hat{\sigma}^2_{\text{jack}}\left(T_n\right) = n\hat{\sigma}^2_{\text{jack}}\left(M_n\right) + \mathcal{O}_p\left(\ell^{\frac{3}{2}} n^{-1}\right). \tag{12}$$

Proof

$$\begin{aligned}\left|n\hat{\sigma}^2_{\text{jack}}(T_n) - n\hat{\sigma}^2_{\text{jack}}\left(M_n\right)\right| &= \left|n\hat{\sigma}^2_{\text{jack}}\left(S_n\right) + 2\frac{(n-\ell)^2}{\ell m}\sum_{j=1}^{m}\left(M_n^{(j)} - M_n^{\cdot}\right)\left(S_n^{(j)} - S_n^{\cdot}\right)\right| \\ &\leq n\hat{\sigma}^2_{\text{jack}}\left(S_n\right) + 2n\left(\hat{\sigma}^2_{\text{jack}}\left(S_n\right)\hat{\sigma}^2_{\text{jack}}\left(M_n\right)\right)^{\frac{1}{2}} \\ &\leq \frac{(n-\ell)^2}{\ell m}\sum S_n^{(j)2} + 2n\left(\hat{\sigma}^2_{\text{jack}}\left(S_n\right)\hat{\sigma}^2_{\text{jack}}\left(M_n\right)\right)^{\frac{1}{2}} \\ &= \mathcal{O}_p\left(\ell^3 n^{-2}\right) + \mathcal{O}_p\left(\ell^{\frac{3}{2}} n^{-1}\right).\end{aligned} \tag{13}$$

■

The sufficient conditions for ***L*** (S';,))2 = *('JP (t⁴n-³)* for the different kinds of statistics are given in the lemma 4.1--4.3of Kunsch (1989).Then for the difference between *LX)* and *M* \ assumption (C) in Kunsch (1989) describes the order of the difference between n&fack *(Mn)* and n&fack *(Ln)* that

$$n\hat{\sigma}^2_{\text{jack}}\left(M_n\right) = n\hat{\sigma}^2_{\text{jack}}\left(L_n\right) + \mathcal{O}_p\left(n^{-\frac{1}{2}}\right) + \mathcal{O}_p\left(\ell n^{-1}\right). \tag{14}$$

Then we have a natural consequence.

Theorem S. Under the former conditions, we have

$$n\hat{\sigma}^2_{\text{jack}}(T_n) = n\hat{\sigma}^2_{\text{jack}}(L_n) + \mathcal{O}_p(\max(\ell^{\frac{3}{2}} n^{-1}, n^{-\frac{1}{2}})). \tag{15}$$

DOI: 10.1111/jtsa.12714

We can see that when -t' - oo and -t' = o (n½), the nonlinear part will decay and linear part will dominate all effects. Under the conditions on the /F(X, F), the *no-Jack* (*Ln*) will converge in probability to the asymptotic variance of *Tn*.

5. SIMULATION

Here, we use some numerical examples to illustrate what we have developed in the former sections. We apply our stationary jackknife variance estimator in different scenarios. Also, we will compare our method stationary jackknife (SJ) with several existing resampling variance estimator such as Moving Block Bootstrap (MBB), Stationary Bootstrap (SB) and Moving Block Jackknife (MBJ).

5.1. Two simple models

In this part. we follow the examples in section 3 of Politis and Romano (1994). The simulated data comes from two MA models

$$X_t = Z_t + Z_{t-1} + Z_{t-2} + Z_{t-3}, \tag{Model 1}$$

and

$$X_t = Z_t - Z_{t-1} + Z_{t-2} - Z_{t-3} + Z_{t-4}, \tag{Model 2}$$

where $Z_1 \overset{iid}{\sim}$ N(0,1). We generate $N = 1000$ observations X_l from each model and we want to estimate the quantity Var(yNX1) by the stationary jackknife and the moving block jackknife. Moreover, we are interested in investigating the estimation performance under a wide range of the (expected) block length. From the derivation in Section 3 we know that the optimal block length for the case of the scaled variance of the arithmetic mean is of order *N*,. In our case this quantity is 1000, = 10. In the simulation, we set the range of the block length as -t' E [1, 80] which contains the optimal block length for both stationary jackknife and moving block jackknife as well as some extreme block length. In the stationary jackknife the corresponding parameter *pin* the truncated geometric distribution is – In Model 1 we have *r: xt 41;: zp* thus *Var(VWOOX1)* 16Var(v-1u\NZ1) = 16. Figure 1 shows the result under Model 1. In Model 2 we have $\sum_{Lr=I}^{1000} X_l$ *Lt* = *I* Z1, so we expect that *Var* (y!OOOX1) var(y'iooo.z1) = 1. Figure 2 shows the result under Model 2. The main difference between the two models is that the autocovariance in Model 1 is always positive for all lags *k* E [0,3] while the autocovariance alternates in sign until the lag *k* is greater than 4 in Model 2. In both figures, we use sj to represent the stationary jackknife and mbj to represent the moving block jackknife. From the results, we can see (especially from the second graph) the moving block jackknife estimator is more sensitive to the block length and the estimation is not reliable when the block length is not chosen suitably. The stationary jackknife gives a reasonable estimation over a wide range of the block lengths. This gives us an intuition that when we do not know the optimal block length for the block jackknife beforehand, the stationary jackknife estimator can be the better choice for inference.

5.2. Application in time series model

Here, we are interested in estimating the variance of sample mean of different time series data with the optimal block length order. In each model we will use all four resampling methods to estimate the variance. In each model we will generate $N = 1000$ data as our original data X1. The block length is an important variable in all methods. Therefore, we will use different values of block length -t' E {*N*,, 2*N*,, 3*N*,} in our simulation.

DOI: 10.1111/jtsa.12714

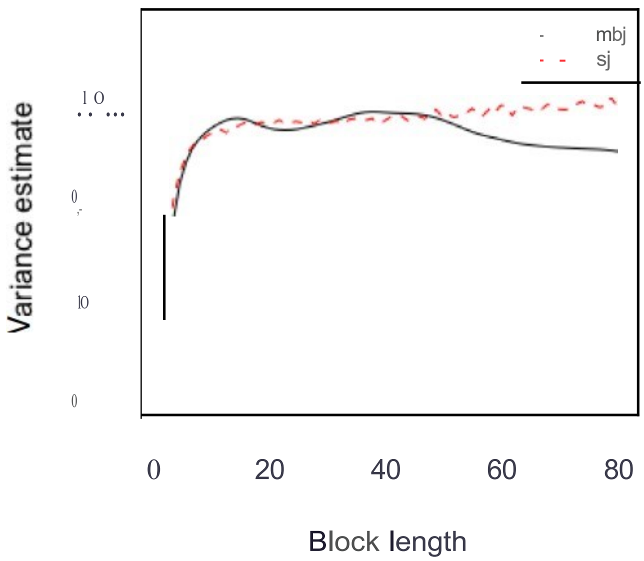

Figure 1. The stationary jackknife and Block Jackknife variance estimator for model $X_t = Z_t + Z_{t-1} + Z_{t-2} + Z_{t-3}$

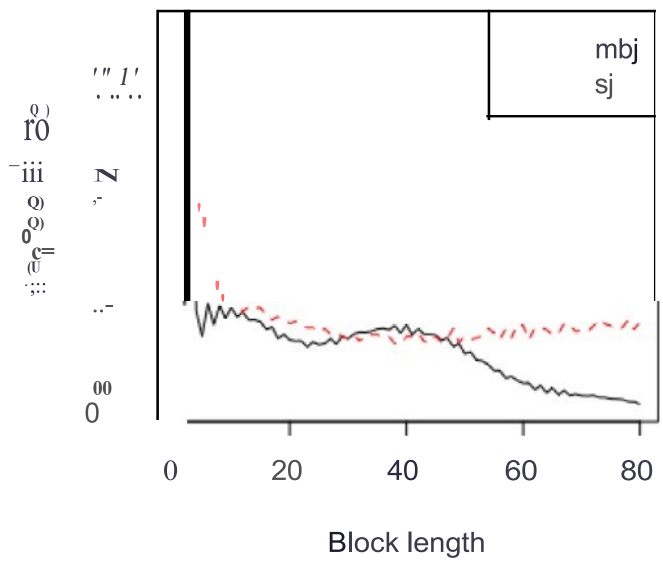

Figure 2. The stationary jackknife and the Block Jackknife variance estimator for model $X_t = Z_t - Z_{t-1} + Z_{t-2} - Z_{t-3} + Z_{t-4}$

The variance estimator is a level-2 estimator. The level-2 estimator (Athreya and Lahiri, 2006) is the estimator which relies on the sampling distribution of a level-1 estimator. The level-1 estimator directly relies on the sampling distribution of the observations. **MSE** of an estimator is a common level-2 estimator. The variance can be treated as the MSE of sample mean. For MBB and SB, we need a positive number B as the number of bootstrap replicates to calculate the Monte Carlo approximation of MBB and SB estimator for this level-2 parameter. We choose $B = 1000$ in our situation. Lastly, to get the estimator of SE of variance estimator, we do 800 simulations for each model.

DOI: 10.1111/jtsa.12714

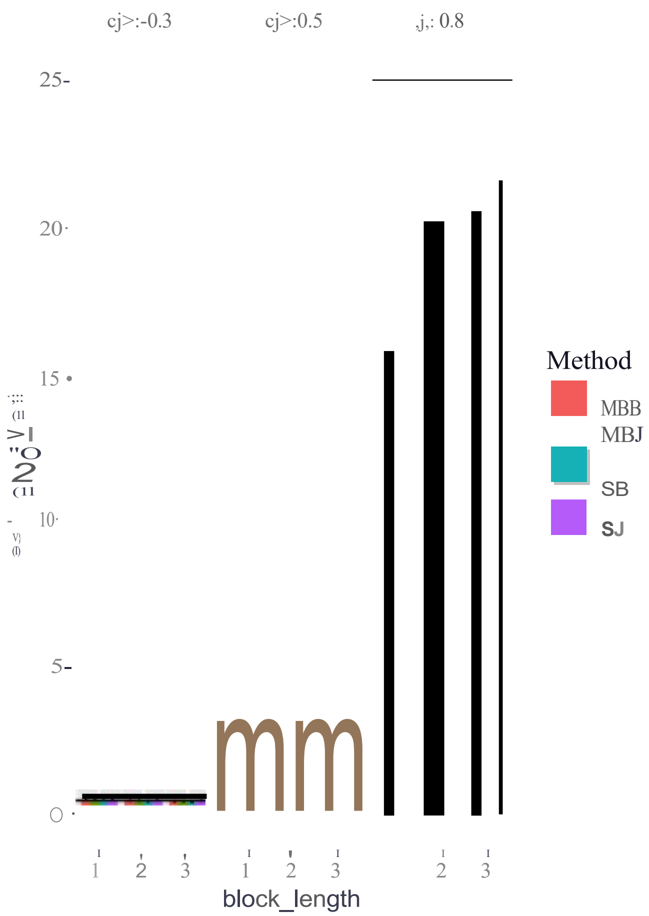

Figure 3. Comparison of four resampling methods in the AR processes

5.2.1. AR(]) model

The first example follows the AR(l) model $X_t = $ </JX,_1 $+ w_t$ in section 3 of Politis and Romano (1994), where $w_t \overset{iid}{\sim} N(0, 1)$. We consider the AR parameter $<p$ vary in {-0.3,0.5,0.8}. We want to estimate the standardized variance of the mean of the data. &$^2 = Var(\sqrt{N}\bar{X}_N)$ Figure 3 shows the results of three different AR(1) processes. The black solid line in each facet is the true variance of the arithmetic mean. We use the coefficients {1, 2, 3} to represent thecorresponding blocklength $\{N_l^1, 2N_l^1, 3N_l^1\}$.The black interval oneach barrepresents theconfidence interval of the estimation. Moreover, we record the computing time for four methods in Figure 4.

5.2.2. *ARMA(2,I) model*

The second modelt́s a ARMA(2,1) model $X_t - 0.6X_{t-1} + 0.05X_{t-2} = W_t + 0.2W_{t-1}$, where $w_t \sim N(0, 1)$. It t́s a causal and invertible ARMA processes. This model can be expressed as $(1 - 0.5B)(1 - 0.1B)X_t = (1 + 0.2B)W$, where B is the backward shift operator. The $Var(\sqrt{N}\bar{X}_N)$ converges to a finite quantity which means *u s* is well-defined. However the computation of this value is a little bit complex, so we use the standardized standard variance of the mean value of data from the 200 simulations to approximate the true value. Figure 5 shows the result under the ARMA model. The red solid line represents the true variance 6.994.

DOI: 10.1111/jtsa.12714

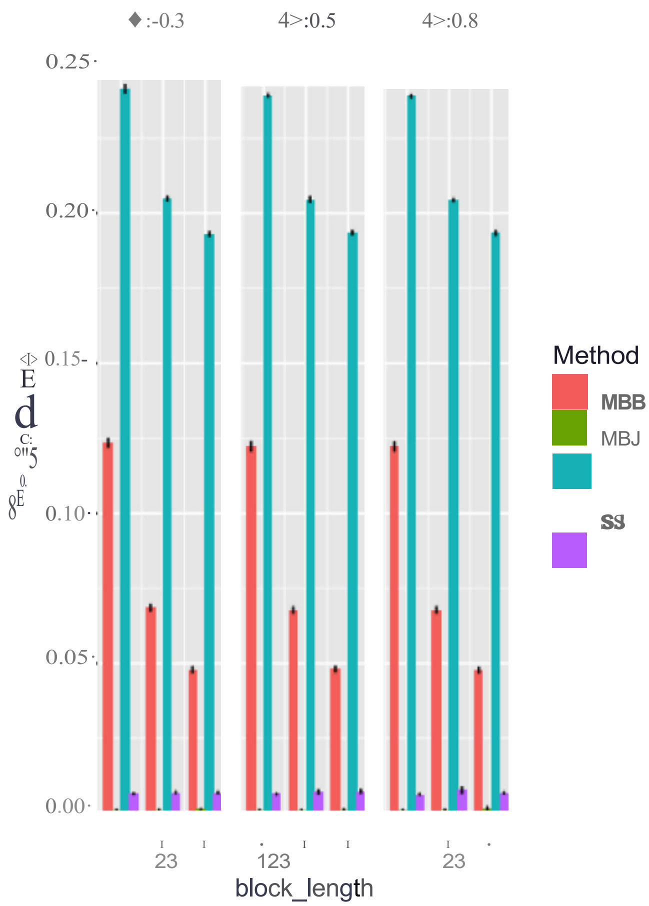

Figure 4. Comparison of four resampling methods' computing time in the AR processes

From the comparison of the four methods in estimating the standardized variance of the mean in different time series methods, the behavior of four methods are pretty similar and all four methods give a reasonable estimate when block length is $3N^{\frac{1}{3}}$. The block jackknife and the stationary jackknife give us the variance estimator with the smaller bias. In the AR(1) process, there is a gap between estimation from all four methods and the true value when $\phi = 0.8$. This is due to the fact that the true value is computed as the asymptotic variance. We have a limited sample size 1000 and the large **AR(1)** coefficient makes the covariance decay much slower than the cases when AR(1) coefficient is 0.3.

Since variance estimator is a level-2 parameter, the bootstrap method needs the Monte Carlo approximation which requires more computation compared with the jackknife method. This can be clearly found in Figure 4. Also, both the stationary jackknife and the stationary bootstrap need to generate random samples from a geometrical distribution. They will consume more time compared with the corresponding moving block methods.

5.3. Simulation for **AR(1)** parameter

The AR(1) model $X_t = \phi X_{t-1} + \epsilon_t$, where $\epsilon_t \overset{iid}{\sim} N(0,1)$. Denote $\hat{\phi}$ as the least square estimator of ϕ, and this is a nonlinear statistic of $\{X_t\}$. We investigate the behavior of stationary jackknife estimator on the variance

DOI: 10.1111/jtsa.12714

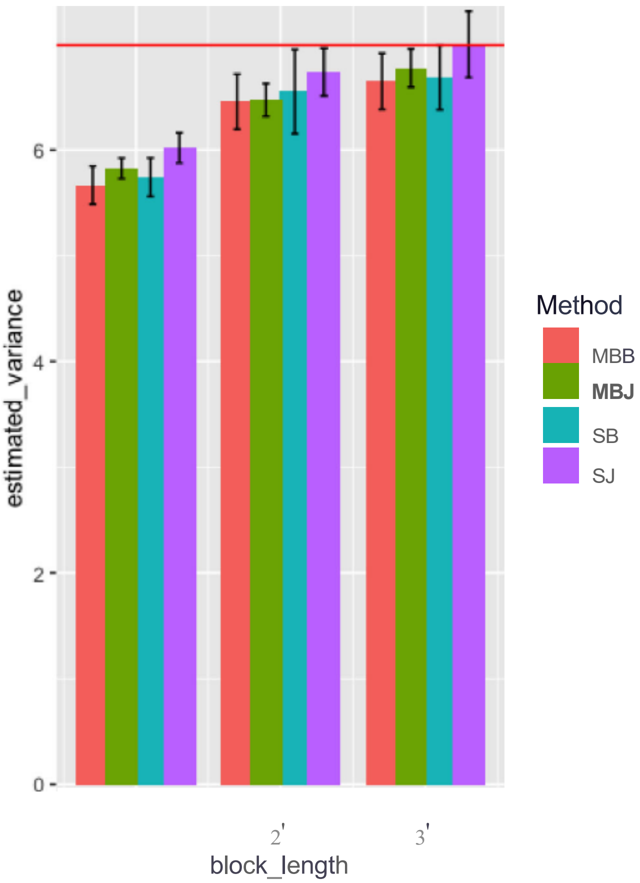

Figure 5. Comparison of four resampling methods in the ARMA(2,I) processes

Table I. The variance estimate by stationary jackknife for AR(]) parameter with different block length (SE is in parenthesis).

	$bl=2$	$bl=3$	$bl=4$	$bl=5$
SJ	0.791(0.013)	0.732(0.014)	0.723(0.014)	0.743(0.0 l6)
PB	0.786(0.015)	0.724(0.013)	0.719(0.012)	0.731(0.013)
True	0.75	0.75	0.75	0.75

estimation of ϕ. Based on the theoretical result, we have yn(<P- ϕ) N(O,1 - ϕ^2). In the simulation setting, $\phi = 0.5$, data length is 200 and the block length l varies in $\{2,3,4,5\}$. For each block length, we repeat the simulation for 500times to compute the SE of the estimator. For reference, we also compute the variance estimator produced by the pair-wise bootstrap (PB) Efron and Gong (1983). The number of bootstrap replication is $B = 500$ in this simulation. Result is in Table I.

From the result, we can see thatthe stationary jackknifegives the reasonable estimators to the variance of **AR(l)** parameter. The behavior is pretty similar to the pairwise bootstrap.

DOI: 10.1111/jtsa.12714

5.4. Trimmed mean

In this part, we check the stationary jackknife variance estimator's performance on the trimmed mean of an AR process. The trimmed mean is a method of averaging after removing the designated percentage of the smallest and the largest values. In our case, we focus on the 60% trimmed mean. This means that we need to remove the 20% extreme values on both ends of the data beforehand. Formally, for the observations $\{X_i\}_{i=1}^{n}$ it is defined as

$$\mu_{tm} = \frac{1}{[0.6n]} \sum_{i=[0.2n]}^{[0.8n]} X_{(i)},$$

where $[a]$ is the largest integer not exceeding a and $X_{(i)}$ is the ith-order statistics of $\{X_i\}$.

In the simulation, the quantity we are interested in is the variance of the scaled mean $\mathrm{var}(\sqrt{[0.6n]}\,\mu_{tm})$. We investigate the stationary jackknife estimator performance for the trimmed mean in different AR(1) models. We choose the AR(1) coefficient α in [0.2,0.5,0.8] and in each AR(1) model we generate $n = 1000$ samples. In each setting, we repeat the experiment for 800 times to get the expectation and the SE. For the different strength of the correlation, we choose the different expected block length. The block lengths are 9, 18 and 27 for

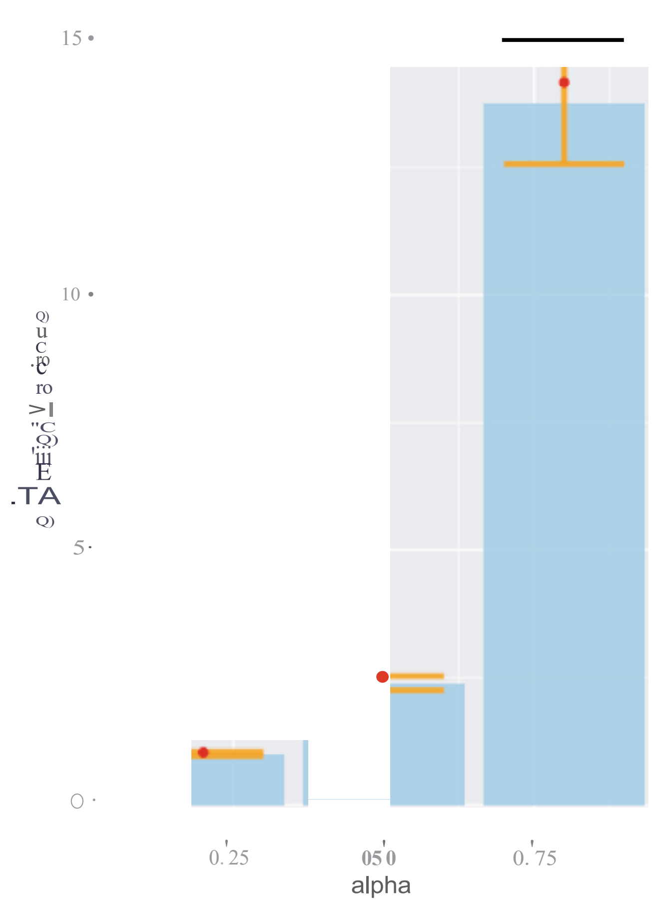

Figure 6. The stationary jackknife estimation on the trimmed mean

DOI: 10.1111/jtsa.12714

the cases when AR(1) coefficients are 0.2, 0.5, and 0.8. The block lengths are approximate to $n^{1/3}$, $2n^{1/3}$ and $3n^{1/3}$. The true variance of the trimmed mean is approximated by the Monte Carlo method. The number of replications in the Monte Carlo method is 10,000. For the convenience of the visualization, we scale both the estimated value and the true value by the sample size. Figure 6 shows the result of the stationary jackknife estimation. The three blue bars represent the value of the stationary jackknife estimator, the orange intervals represent the confidence interval for the estimation and the red points are the true variance under three different data generating processes.

From the result, we can see the stationary jackknife variance estimation is pretty close to the true value. Moreover, the true variance is inside the confidence interval of the stationary jackknife estimation.

5.5. Personal savings rate data

In this section, we compare the four resampling methods on personal saving rate (psavert) data. Psavert is calculated as the ratio of personal saving to disposable personal income (DPI). The rate can be generally viewed as the portion of personal income that is used for investments. In Figure 7, we have monthly psavert data in a 30-year interval from September 1988 to August 2018 (totally 360 data points). We take the first difference of the series

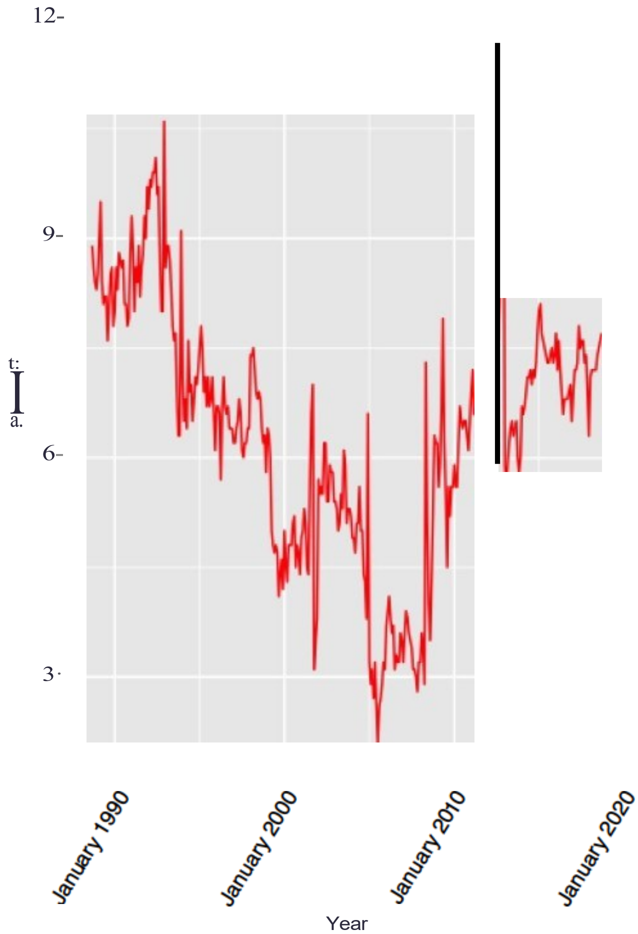

Figure 7. Monthly personal saving rate (psavert) time series in 1988-2018

DOI: 10.1111/jtsa.12714

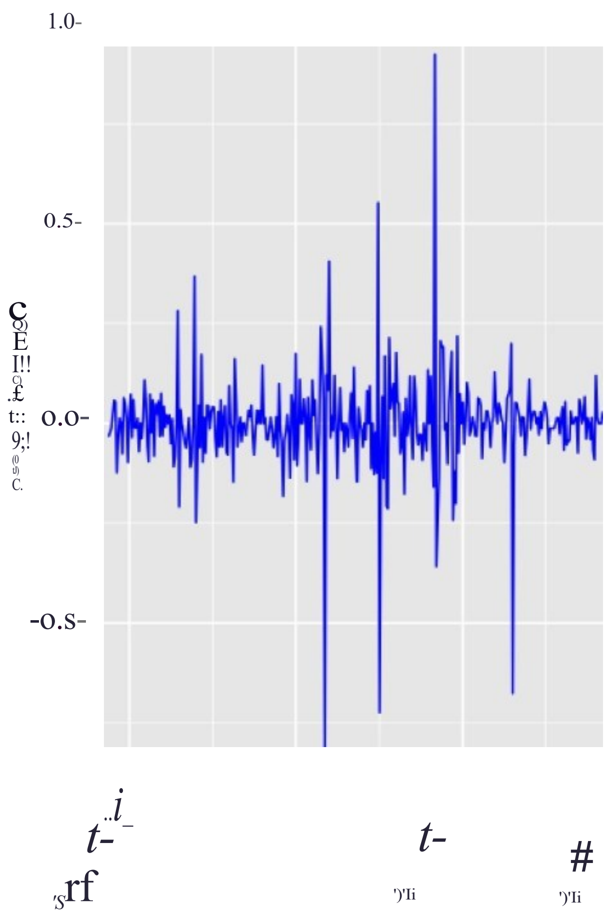

Figure 8. Monthly personal saving rate time series (psavert) increment in 1988-2018

Table II. The SE of the ARMA(l,1) coefficients under *four* methods

	MBB	SB		SJ
AR_coefficient	0.135	0.141	0.138	0.140
MA_coefficient	0.114	0.120	0.112	0.116

to make it stationary. The transformed series (Figure 8) can be interpreted as the personal saving rate increment. An ARMA(1,1) model is used to model psavert increment series $\{Y_i\}$,

$$Y_t = 0.374Y_{t-1} + \epsilon_t - 0.756\epsilon_{t-1}.$$

We are interested in the performance of the four resampling methods in estimating the SE of the coefficients in the ARMA(l,l) model. For simplicity, we introduce the AR_coefficient and MA_coefficient to represent the corresponding coefficients. In this case, we adopt $t' = 8$ ($p = 1/8$) that is roughly equal to N^{1}. For Bootstrap methods (MBB and SB), we choose $B = 800$. The result is shown in Table II. The SB and SJ have slightly larger

DOI: 10.1111/jtsa.12714

estimates which is coincidence with the randomness in the method. Moreover, a few extreme values also amplify the SB SE estimator.

6. CONCLUSION

The main contribution of this article is to propose a new resampling method, called the stationary jackknife, for estimating the variance of a statistic of interest for weakly dependent data. Consistency of variance estimation is proved under certain assumptions for both the sample mean and for certain nonlinear statistics. The results on the optimal block length order are consistent with the existing results for the moving block jackknife. From both simulation data and the real data, it is evident that the stationary jackknife estimator is comparable to other classical resampling methods. Because of the simple form of the stationary jackknife, the computing time is much lower when compared to Bootstrap methods. Moreover, the stationary jackknife can provide reasonable estimates for a wide range of the (expected) block lengths, primarily due to the introduction of the randomness in the blocklengths.

ACKNOWLEDGEMENTS

This research is partially supported by NSF grants No: DMS 2006475, 2131233 and 2235457.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

REFERENCES

Athreya K, Lahiri S. 2006. *Measure Theory and Probability Theory.* Springer, New York.

Carlstein E. 1986. The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *The Annals of Statistics* 14(3):1171-1179.

Doukhan P. 1994. *Mixing: Properties and Examples,* Vol. 85. Springer-Verlag, New York.

Efron B. 1979. Bootstrap methods: another look at the jackknife. *The Annals of Statistics* 7:1-26.

Efron B, Gong G. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. *The American Statistician* 37(1):36-48.

Hoeffding W. 1948. A class of statistics with asymptotically normal distributions. *Annals of Statistics* 19:293-325.

Kunsch **H.** 1989. The jackknife and the bootstrap for general stationary observation. *The Annals of Statistics* 17:1217-1241.

Lahiri SN. 1999. Theoretical comparisons of block bootstrap methods. *Annals of Statistics* 27(1):386-404.

Lahiri SN. 2003. *Resampling Methods for Dependent Data.* Springer Science & Business Media, New York.

Miller R. 1974. The jackknife- a review. *Biometrika* 61:1-15.

Nordman D. 2009. A note on the stationary bootstrap's variance. *The Annals of Statistics* 37(1):359-370.

Park D, Willemain T. 1999. The threshold bootstrap and threshold jackknife. *Computational Statistics & Data Analysis* 31(2):1171-1179.

Pellegrino, F. (2022). Selecting time-series hyperparameterswith the artificial jackknife. *arXiv preprint arXiv:2002.04697.*

Politis D, Romano J. 1992. A circular block-resampling procedure for stationary data. *Exploring the Limits of Bootstrap* (pp. 263-270). Wiley, New York.

Politis D, Romano **J.** 1994. The stationary bootstrap. *Journal of the American Statistical Association* 89(428):1303-1313.

Quenouille **M.** 1949. Approximate tests of correlation in time-series. *Journal of the Royal Statistical Society, Series B (Methodological)* 11(1):68-84.

Quenouille M. 1956. Notes on bias in estimation. *Biometrika* 43:353-360.

Shao J, Wu C. 1989. A general theory for jackknife variance estimation. *The Annals of Statistics* 17(3):1176-1197.

Tukey J. 1958. Bias and confidence in not quite large samples(abstract). *The Annals of Mathematical Statistics* 29(614).

Wu C. 1986. Jackknife, bootstrap and other resampling methods in regression analysis. *The Annals of Statistics* 14(4):1261-1295.

DOI: 10.1111/jtsa.12714

APPENDIX A

A.1. *Proof of Theorem 1*

Proof By the simple algebraic operations, for any constantcthejth stationary jackknife replication can be expressed as

$$T_n^{(j)} = \bar{X}_n + \frac{1}{n - \tilde{L}_j}\left[\tilde{L}_j\left(\bar{X}_n - c\right) - \sum_{t=j}^{j+\tilde{L}_j-1}\left(X_t - c\right)\right]. \tag{A1}$$

We let

$$c = \hat{\mu}_n = \frac{\sum_{j=1}^{m}\frac{1}{n-\tilde{L}_j}S_{j,\tilde{L}_j}}{\sum_{j=1}^{m}\frac{\tilde{L}_j}{n-\tilde{L}_j}},$$

where $S_{j,\tilde{L}_j} = \sum_{t=j}^{j+\tilde{L}_j-1} X_t$.
Now we have

r,·=) n _!_ m l III L,,J '[V) II ·

$$= \bar{X}_n + \frac{1}{m}\sum_{j=1}^{m}\frac{1}{n-\tilde{L}_j}\left(\tilde{L}_j\left(\bar{X}_n - \hat{\mu}_n\right)\right).$$

Moreover

TV)_ r,·=> _ tf. — _!_ 111 L.J ;L_) X n n (n - Lj m i- 1 n - L;

-(n S- \-¾B S n \i · (A2)

After plugging the formula (A2) of T/ - 7;.·>into the expression of no-Jack' we have

na-2= Jack (n- t')2 mt' 111 [(£..J .F 1 L-J - n - L- J .!. m III £..J ... 1 L-J n - £. J) X- n (S.0L n-' £. J ..!. m 111 £. J S ,n- £ .)] 2

= 1+ t) (t1og (n)) L mt'[1- 1 (L 2- n L m 2- m- 1 n L) 2 X ! L m S2 j,Lj n - ! m (L m S j,Lj) 2 n

- 2L 1 a. 2.- ! 11 L L n n n (S j,L; n - L S± m is=t n | xn . (A3)

DOI: 10.1111/jtsa.12714

Notice that the second equality is due to the following

$$\frac{1}{(n-\tilde{L}_i)^2} = \frac{n^2}{(n-\tilde{L}_i)^2}\frac{1}{n^2}$$
$$\le \frac{n^2}{(n-2\ell\log n)^2}\frac{1}{n^2}$$
$$= \left(1+\mathcal{O}\left(\frac{\ell\log n}{n}\right)\right)\frac{1}{n^2}.$$

Under the Assumption 1, we have ${}_{n}$ 0. This means

$$\mathcal{O}\left(\frac{\ell\log n}{n}\right) = o\left(\frac{\ell}{n^{\frac{2}{3}}}\right)$$
$$=o(1).$$

So we only consider the leading term in (A3).

$$n\hat{\sigma}^2_{\text{jack}} = \frac{n^2}{ml}\left[\sum_{j=1}^{m}\left(\frac{\tilde{L}_j}{n}-\frac{1}{m}\sum_{i=1}^{m}\frac{\tilde{L}_j}{n}\right)^2\bar{X}_n^2 - \frac{1}{m}\left(\sum_{j=1}^{m}\frac{S_{j,L_j}}{n}\right)^2 + \sum_{j=1}^{m}\frac{S^2_{j,L_j}}{n^2}\right.$$
$$\left. - 2\sum_{j=1}^{n}\left(\frac{\tilde{L}_j}{n}-\frac{1}{m}\sum_{i=1}^{m}\frac{\tilde{L}_i}{n}\right)\left(\frac{S_{j,L_j}}{n}-\frac{1}{m}\sum_{i=1}^{m}\frac{S_{i,L_i}}{n}\right)\bar{X}_n\right]$$
$$= A_1 - A_2 + A_3 - A_4.$$

We decompose the leading term of n&!ck into four parts and investigate the order of the each part.

For A_1,We have

$$\mathbb{E}(A_1) = \mathbb{E}\frac{n^2}{\ell m}\sum_{j=1}^{m}\left(\frac{\tilde{L}_j}{n}-\frac{1}{m}\sum_{i=1}^{m}\frac{\tilde{L}_i}{n}\right)^2\bar{X}_n^2$$
$$= \frac{n}{ml}\mathbb{E}n\bar{X}_n^2\mathbb{E}\left(\sum_{j=1}^{m}(\frac{\tilde{L}_j}{n}-\frac{1}{m}\sum_{i=1}^{m}\frac{\tilde{L}_i}{n})^2\right)$$
$$= \frac{n}{ml}\mathbb{E}n\bar{X}_n^2\mathbb{E}\left(\sum_{j=1}^{m}\left(\frac{\tilde{L}_j^2}{n^2}\right) - \mathbb{E}\frac{1}{m}\left(\sum_{i=1}^{m}\frac{\tilde{L}_i}{n}\right)^2\right)$$
$$= \frac{n}{ml}\left(\sum_{k=-\infty}^{\infty}|R(k)|\right)\mathbb{E}\left(\sum_{j=1}^{m}\frac{\tilde{L}_j^2}{n^2}\right)$$
$$= \mathcal{O}\left(\frac{\ell}{n}\right).$$

DOI: 10.1111/jtsa.12714

During the derivation, we use the following two facts,

$$\mathbb{E}n\bar{X}_n^{\,2} = \sum_{k=-n+1}^{n-1}\left(1-\frac{|k|}{n}\right)R(k) \leq \sum_{k=-n+1}^{n-1}|R(k)| \leq \sum_{k=-\infty}^{\infty}|R(k)|.$$

and

$$\mathbb{E}(L_j^2) = \sum_{i=1}^{[2t'\log n]-1} i^2 p(1-p)^{i-1} + [2t'\log n]^2(1-p)^{[2t'\log n]-1} \leq \sum_{i=1}^{\infty} i^2 p(1-p)^{i-1} = \frac{2-p}{p^2}.$$

For part A_2, we consider the coefficient of the X_t in the summation $\sum_{i=1}^m S_{i,L_i}$. We denote

$$\sum_{i=1}^{m} S_{i,L_i} = \sum_{t=1}^{n} b_t X_t,$$

where b_t is a random variable with the following relationship

$$b_t = \sum_{j=1}^{t} \mathbb{1}(L_j \geq t-j).$$

Suppose $\mathcal{L} = \sigma(L_1, L_2, \ldots, L_m)$, the sigma algebra generated by the block length variables. Hence $\mathbb{E}\left(\sum_{j=1}^m S_{j,L_j}\right)^2$ can be expressed as

$$\mathbb{E}\left(\left(\sum_{j=1}^{m} S_{j,L_j}\right)^2\right) = \mathbb{E}\left(\mathbb{E}\left(\left(\sum_{j=1}^{m} S_{j,L_j}\right)^2 \mid \mathcal{L}\right)\right) = \mathbb{E}\sum_{k=-n+1}^{n-1} v(|k|)R(k),$$

where $v(|k|) = \sum_{t=1}^{n-|k|} b_t b_{t+k}$.

Now we focus on the quantity $\mathbb{E}(b_t b_{t+k})$. For any $k \geq 0$, we have

$$\mathbb{E}b_t b_{t+k} = \mathbb{E}\sum_{j=1}^{t+k}\mathbb{1}(L_j > t+k-j)\sum_{j=1}^{t}\mathbb{1}(L_j > t-j)$$
$$\leq \mathbb{E}\sum_{j=1}^{t+k}\mathbb{1}(L_j > t+k-j) + \mathbb{E}\sum_{j=1}^{t+k}\mathbb{1}(L_j > t+k-j)\,\mathbb{E}\sum_{j=1}^{t}\mathbb{1}(L_j > t-j)$$

DOI: 10.1111/jtsa.12714

t+k t+k I
_L/1 - p)'+k-j + _L/1- py+k-j.LP - py-i
j=l j=l j=l

<-1+-1
- p p2"

Consequently, we have

$$v(|k|) \leq \frac{n-|k|}{p} + \frac{n-|k|}{p^2}.$$

Based on the bound ofv(lkl), we can derive the bound for A_2,

IIE(A2)I+-2__Li n-1 (t'+t'2) IR(k)I
m nt' k=-n+t

=! IR(k)I+o(!)

=C?(;)·

For partA_3

$$\mathbb{E}(A_3) = \frac{n}{m\ell}\sum_{j=1}^{m}\mathbb{E}\left(\mathbb{E}\left(\frac{S^2_{j,L_j}}{n}|\tilde{L}_j\right)\right)$$

$$= \frac{n}{\ell}\mathbb{E}\left(\mathbb{E}\left(\frac{S^2_{j,L_j}}{n}|\tilde{L}_j\right)\right)$$

= [2t"togn]-t s-l
L L (s-lkl)R(k)p(l-p)'_l
/ (S=l k=-s+t n

+ [ui]-t ([2t'logn]n-lkl)R(k)(l-p)[2t"togn]l
k=-[2t"logn]+l

=i [Ulogn]-l [2t"logn]
_Li _Li (s- lkl)R(k)p(l - p)s-l +(? c)
k=-[2t"togn]+l S=lkl+l

=i [Ulogn]-l [2t"logn]-lkl
L L (1-p)'k'R(k)tp(l-p)'-1 +C?(n)
k=-[2t"logn]+l 1=1

= [Ulogn]-l
L (1 - p)lkl+l - (1- Pi2t"logn]R(k) + (? (:2) +(?(n)
k=-[2t"togn]+t

s r [2t'logn]-I k~-[2t'logn]+I (1 +p(lkl + l))R(k) + (')C)

= CT + (')(7) .

The derivation is based on repeatedly using the Lemma 2.

Lemma 2. Assume that p +(npr¹ 0 as n oo. Let tn be a sequence of positive number such that tn oo as n oo. L follows a Geometric distribution with the parameter p. Then we have

$$P\left(L_1 > t_n p^{-1}\right) = \mathcal{O}\left(exp\left(-t_n\right)\right).$$

Proof Let $q = 1 - p$, then we have

$$P\left(L_1 > t_n p^{-1}\right) = \sum_{i > p^{-1} t_n} pq^{i-1}$$

s qp !a.-t

= exp([p-\- 1]Iogq)

= (')(exp(-p (p-Itn - 1)))

= (')(exp (-tn)) .

■

For the cross part A4, from the result of the former part it follows from the Cauchy-Schwarz inequality that

$$\mathbb{E}\left(A_4\right) = \mathbb{E}\left(\frac{2n^2}{m\ell}\sum_{j=1}^{n}\left(\frac{\tilde{L}_j}{n} - \frac{1}{m}\sum_{i=1}^{m}\frac{\tilde{L}_i}{n}\right)\left(\frac{S_{j,L_j}}{n} - \frac{1}{m}\sum_{i=1}^{m}\frac{S_{i,L_i}}{n}\right)\bar{X}_n\right)$$

$$\leq \frac{n^2}{m\ell}\left(\mathbb{E}\sum_{j=1}^{m}\left(\frac{\tilde{L}_j}{n} - \frac{1}{m}\sum_{i=1}^{m}\frac{\tilde{L}_i}{n}\right)^2 \bar{X}_n^2\right)^{\frac{1}{2}}$$

$$\left(\mathbb{E}\sum_{j=1}^{m}\frac{S_{j,L_j}^2}{n^2} - \frac{1}{m}\left(\sum_{j=1}^{m}\frac{S_{j,L_j}}{n}\right)^2\right)^{\frac{1}{2}}$$

$$= \mathcal{O}\left(\left(\frac{\ell}{n}\right)^{\frac{1}{2}}\right).$$

Finally based on the results of the A A A and A we have

lE(n&I,ck) =E(A1 +A2+A3+A4),

=u;.+(')(7)+(')(( )½)-

■

A.2. Proof of Theorem 2

Proof Without loss of generality, we assume IE (X;) = 0. Then it is obvious that

$$
\begin{aligned}
\text{IE}(Sp,;) &= \text{IE}(\ \text{IE}(\ SiL;|Li)) \\
&= \text{IE}\,(Li)\,\text{IE}\,(Xi) \\
&= 0.
\end{aligned}
$$

We look at L"$_{1\text{-}l}$Cov^2 (s$_{l}$r$_{l}$,S,-r), it can be decomposed into $\sum_{j=1}^{[2\ell \log n]} \mathrm{Cov}^2(S_{1,\tilde{L}_1}, S_{j,\tilde{L}_j})$ and L_;:[Ulogn]+lCov^2 (sI,r,,SiL;)· Whenj [2t'logn] +1, wehave

$$\mathrm{Cov}\left(S_{1,\tilde{L}_1}, S_{j,\tilde{L}_j}\right) = \mathbb{E}(S_{1,\tilde{L}_1} S_{j,\tilde{L}_j}) = \mathbb{E}(\mathbb{E}(S_{1,\tilde{L}_1} S_{j,\tilde{L}_j} | \tilde{L}_1, \tilde{L}_j))$$

= [Ulogn]-l Li k-l IESI,ksj,Lf(l - Pi-l +IESI,[Ulogn]sj,L/1 - p)IUlogn]-l

= ([Ulogn]-[/!] Li k-1 + [Ulogn]-1 Li [Ulogn]- [t'/] +I IES1,ksj.rf(l - Pi-1

+(') t'(logn)½) / n2

S [U logn]-[t' I ''' Li a 12!2. (i- k)IISI.kIIIiE. IIS,-r.I1 6+/J/2 ' / 6+FJ/2 p(I –p)"-l

$$+ \sum_{k=[2\ell \log n]-[\ell^{\frac{1}{2}}]+1}^{[2\ell \log n]-1} \sqrt{\mathbb{E}S_{1,k}^2 \mathbb{E}S_{j,\tilde{L}_j}^2}\, p(1-p)^{k-1} + \mathcal{O}\left(\frac{\ell \log^{\frac{1}{2}} n}{n^2}\right)$$

S a 12!2o (i - [2t'log n] +[t'½J)CIpt'½ [Ulogn)-[t'!] Li (k-l k(l - p?k.-lJl ½

$$\left([2\ell \log n] - \left[\ell^{\frac{1}{2}}\right]\right)^{\frac{1}{2}} + \mathcal{O}\left(\frac{\ell (\log n)^{\frac{1}{2}}}{n^2}\right)$$

= (') (tlogin) aii+zi(i- [2t'logn] + [t'2]) + (') (t'(logn) / n2).

DOI: 10.1111/jtsa.12714

Hence we have

$$\sum_{j-[Ulogn]+t}^{m} \text{Cov}^2(S_{1,r,},S_{j,LJ} \quad \sum_{fa-[Ulogn]+t}^{m} \underline{r>(t'^2 logn)a.!.(j-[2t'logn]+[t'½])}$$

$$\sum_{fa-[t! I}^{1} \underline{t'a.!.} (j)C>(t'logn)$$

$$\sum_{fa-[t! I}^{1} \underline{fa.!.} (j)C>(t'logn)$$

$$= C>(t'logn).$$

When j [2t' log $n]$, we can write the expression of Cov ($S_{1,L,}, S_{j.L;})$ as

$$\text{Cov} (S_{1,i,}, s_{j,LJ} = \text{lE} (S_{1,r}s_{j,Z,})$$

$$= \sum_{u--[2t'logn]}^{[Ulogn]} b(u)R(j-l+u),$$

where

$$b(u)= \sum_{i-1}^{[2t'togn]-lul} P(L \ i)P(L \ i+lul)$$

$$= (1 \_ p)lul \sum_{i-1}^{[Ulogn]-lul} (1 \_ p)2<i-1)$$

$$= \frac{\ell}{2}(1-p)^{|u|} - \frac{\ell}{2}(1-p)^{2[2\ell \log n]+2-|u|}$$

$$= \frac{\ell}{2}(1-p)^{|u|} + \mathcal{O}\left(\frac{\ell}{n^4}\right). \quad \text{(A4)}$$

In the last step we use the facts that (1 - p)ll = (′.) (exp (-)) and L is a random variable with the truncated geometric distribution.

From (A4) we know the coefficient for R(O) in Cov (S_{1}.L,,$s_{j,LJ}$ is f(l - p)li-[1]1, which is denoted as at For any j 1, we are interested in the coefficient of R (k), lkl t' in the quantity Cov ($S_{1,i,}, S_{jL;}$)- The reason is that

$$\sum_{|K|>\ell} R(K) = \mathcal{O}\left(\frac{1}{\ell^2}\right).$$

Denote at as the coefficient of $R(k)$ in Cov ($S_{1,L,}, S_{j.L;}$), then for lkl t', we have a bound for a \ that is

$$a_k^{(j)} \in \left[(1-p)^{|k|}a_0^{(j)}, (1-p)^{-|k|}a_0^{(j)}\right].$$

Due to limp o(1 - pf = e-1, we know a%) is the same order of a). Thus we have

$$Cov\left(S_{1,L_1}, S_{j,L_j}\right) = \mathcal{O}\left(a_0^{(j)}\sigma_{as}^2\right). \tag{AS}$$

By repeatedly applying the equation (A4), we have

$$\sum_{j=1}^{[2\ell \log n]} \mathrm{Cov}^2\left(S_{1,L_1}, S_{j,L_j}\right) = \sum_{j=1}^{[2\ell \log n]} a_0^{(j)2}\sigma_{as}^4$$
$$= \frac{\ell^2}{4}\sigma_{as}^4 \sum_{j=1}^{[2\ell \log n]} (1-p)^{2j}$$
$$= \frac{\ell^3}{8}\sigma_{as}^4 (1 + o(1)).$$

Then we need to show that

$$\sum_{j=1}^{m} \left|\mathrm{Cov}\left(S_{1,L_1}^2, S_{j,L_j}^2\right) - 2\mathrm{Cov}^2\left(S_{1,L_1}, S_{j,L_j}\right)\right| = o\left(\ell^3\right). \tag{A6}$$

The key method to prove (A6) is the following inequality, for s S t S u S v,

$$|\mathbb{E}\left(X_sX_tX_uX_v\right) - \mathbb{E}\left(X_sX_t\right)\mathbb{E}\left(X_uX_v\right) - \mathbb{E}\left(X_sX_u\right)\mathbb{E}\left(X_tX_v\right) - \mathbb{E}\left(X_sX_v\right)\mathbb{E}\left(X_tX_u\right)|$$
$$\text{S } Ca(max(t-s, u-t, v-u))^{\bullet}. \tag{A7}$$

This is followed by a repeated application of theorem 17.2.3 of (Politis and Romano, 1994).

In our case, for a certain maximum gap t between variable, there are at most 3k^2[2t' log *n]* sets of indices {*s, t, u,* v} where {*s,* t} E {1, 2 • • • [2t'log *n]}* and {*u,* v} E *U,j* + 1, ... *j* +[2t' log*n]* - 1} for any *j*. Now we have

$$\sum_{j=1}^{m} \left|\mathrm{Cov}(S_{1,L_1}^2, S_{j,L_j}^2) - 2\mathrm{Cov}^2(S_{1,L_1}, S_{j,L_j})\right|$$

S (Li$_{j-1}^{[U' \log n]}$ Li$_{ka:()}^{j+[U' \log n)}$ + Li$_{[U' \log n)+I}^{m}$ Li$_{k-0}^{j+[U'\log n))}$ 3[2t'logn]k2a=(k) •

S Li$_{I}^{[2t'\log n)}$ Li$_{k-0}^{oo}$ L3[2t'logn]k²a (k) + Li$_{j-[2t'\log n)}^{oo}$ Li$_{k-j-[2/t' \log n)}^{j+[U' \log n)}$ 3[2t'logn]k2a•!.(k)

S Li$_{I}^{[2t' \log n)}$ Li$_{1.:-0}^{oo}$ L3[2t'logn]k²a (k) + Li$_{[2t'\log n)}^{oo}$ Li$_{k-j-[U'\log n)}^{j+[U' \log n)}$ 3[2t'logn]k2a.!•(k)

= <9(t'logn) + (Li$_{k-1}^{2[2t'\log n]}$ Li$_{j-[2t'\log n)+I}^{[U'\log n)+k}$ + Li$_{1.:-[U'\log n)+I}^{m}$ Li$_{j-k-[U'\log n)}^{k+[2t'\log n)}$) 3[2t'logn]k²a.!•(k)

= <9(t'logn) + (9 (t^2(logn)2)

$$=_{o} (t^3). \tag{A8}$$

DOI: 10.1111/jtsa.12714

Above derivation is based on the fact that when $j > \lfloor 2\ell \log n \rfloor + 1$ the maximum gap between the indices is at least $j - \lfloor 2\ell \log n \rfloor$ and assumption in theorem 3.2 that $\sum_{k=0}^{\infty} k^2 \alpha^{\delta/(2+\delta)}(k) < \infty$.

Then it is obvious that

$$\sum_{i=1}^{m}\sum_{j=1}^{m} \mathrm{Cov}\left(S_{i,L_i}^2, S_{j,L_j}^2\right) = \sum_{i=1}^{m}\sum_{j=1}^{m} \left(\mathrm{Cov}\left(S_{i,L_i}^2, S_{j,L_j}^2\right) - 2\mathrm{Cov}^2\left(S_{i,L_i}, S_{j,L_j}\right) \right.$$

$$\left. + 2\mathrm{Cov}^2\left(S_{i,L_i}, S_{j,L_j}\right)\right)$$

$$= \frac{m\ell^3}{4}\sigma_{as}^4(1+o(1)). \tag{A9}$$

Back to the variance of our jackknife estimator, we have

$$\mathrm{Var}\left(n\hat{\sigma}^2_{\text{jack}}\right) = \frac{n^4}{m^2\ell^2}\mathrm{Var}\left(\sum_{j=1}^{m}\left(\frac{S_{j,L_j}}{n} + \sum_{i=1}^{n} w_i^{(j)}X_i\right)^2\right)\left(1 + o\left(\frac{\ell}{n}\right)\right),$$

where we denote

$$\left(\frac{\tilde{L}_j}{n} - \frac{1}{m}\sum_{i=1}^{m}\frac{\tilde{L}_i}{n}\right)\bar{X}_n + \frac{1}{m}\sum_{i=1}^{m}\frac{S_{i,L_i}}{n} = \sum_{i=1}^{n} w_i^{(j)}X_i.$$

Notice that $w_i^{(j)}$ are all random variables but it is easily to verify that $|w_i^{(j)}| = \mathcal{O}\left(\frac{2\ell\log n}{n^2}\right)$.

From the derivation above, we can naturally get that

$$\frac{n^4}{m^2\ell^2}\mathrm{Var}\left(\sum_{j=1}^{m}\frac{S_{j,L_j}^2}{n^2}\right) = \frac{1}{m^2\ell^2}\sum_{i=1}^{m}\left(\sum_{j=1}^{n}\mathrm{Cov}^2(S_{i,L_i}, S_{j,L_j}) + o(\ell^3)\right)$$

$$= \frac{1}{m^2\ell^2}\frac{m\ell^3}{4}\sigma_{as}^4(1+o(1))$$

$$= \frac{\ell^2}{4n}\sigma_{as}^4(1+o(1)). \tag{A10}$$

For part $\frac{n^4}{m^2\ell^2}\mathrm{Var}\left(\sum_{j=1}^{m}\left(\sum_{i=1}^{n} w_i^{(j)}X_i\right)^2\right)$, we have

$$\frac{n^4}{m^2\ell^2}\mathrm{Var}\left(\sum_{j=1}^{m}\left(\sum_{i=1}^{n} w_i^{(j)}X_i\right)^2\right) \le \left(\frac{2\ell\log n}{n^2}\right)^4 \frac{n^4}{m\ell^2}\mathbb{E}\left(\sum_{i=1}^{n}X_i\right)^4$$

$$= \mathcal{O}\left(\frac{\ell^2\log^4 n}{n^3}\right)$$

$$= o\left(\frac{\ell}{n}\right). \tag{A11}$$

DOI: 10.1111/jtsa.12714

From this we can seethat in theexpression of the variance of the stationary jackknife variance estimator, the part Var(L_;'. [illegible])plays the dominant role, so we can ignore the other parts and the expression for our variance is

$$\mathrm{var}(n\&^{2}_{\mathrm{Jae}\ \mathrm{k}}) = \frac{!}{4n} \ldots a^{4}_{as} + o\left(\frac{}{n}\right) .$$

■

DOI: 10.1111/jtsa.12714